

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuLink Inc.

OM nucleic acid search using SW model

Run on: January 14 2003 11:42:33 Search time 37.817 seconds

(without alignments)  
11136.303 Million cell updates/sec

Title: US-09-910-428-2

Perfect score: 26

Sequence: 1 cclccccaatcaatcaatcattc 26

Scoring table: IDENTITY\_NTC

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

EST:\*

- 1: em\_estba:\*
- 2: em\_estba:\*
- 3: em\_estba:\*
- 4: em\_estba:\*
- 5: em\_estba:\*
- 6: em\_estba:\*
- 7: em\_estba:\*
- 8: em\_estba:\*
- 9: em\_estba:\*
- 10: em\_estba:\*
- 11: em\_estba:\*
- 12: em\_estba:\*
- 13: em\_estba:\*
- 14: em\_estba:\*
- 15: em\_estba:\*
- 16: em\_estba:\*
- 17: em\_estba:\*
- 18: em\_estba:\*
- 19: em\_estba:\*
- 20: em\_estba:\*
- 21: em\_estba:\*
- 22: em\_estba:\*
- 23: em\_estba:\*
- 24: em\_estba:\*
- 25: em\_estba:\*
- 26: em\_estba:\*
- 27: em\_estba:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21.4	82.3	431	17	BH032631
2	21.4	82.3	431	17	BH032631
3	20.6	83.0	473	17	A415322
4	20.6	79.2	973	17	A415322
5	20.4	78.5	790	17	CNS0700N
6	20.2	77.7	804	17	A415322

Result No.	Score	Query Match	Length	DB ID	Description
7	19.8	76.2	476	19	AM532881
8	19.8	76.2	535	19	AM532881
9	19.8	76.2	537	19	AM532881
10	19.8	76.2	587	19	BH032631
11	19.8	76.2	635	17	A415322
12	19.8	76.2	639	17	A415322
13	19.8	76.2	644	10	AM331729
14	19.8	76.2	670	17	A415322
15	19.8	76.2	683	14	B0199536
16	19.8	76.2	702	17	A415322
17	19.8	76.2	754	17	A415322
18	19.8	76.2	754	17	A415322
19	19.8	76.2	754	17	A415322
20	19.8	76.2	754	17	A415322
21	19.8	76.2	754	17	A415322
22	19.8	76.2	754	17	A415322
23	19.8	76.2	754	17	A415322
24	19.8	76.2	754	17	A415322
25	19.8	76.2	754	17	A415322
26	19.8	76.2	754	17	A415322
27	19.8	76.2	754	17	A415322
28	19.8	76.2	754	17	A415322
29	19.8	76.2	754	17	A415322
30	19.8	76.2	754	17	A415322
31	19.8	76.2	754	17	A415322
32	19.8	76.2	754	17	A415322
33	19.8	76.2	754	17	A415322
34	19.8	76.2	754	17	A415322
35	19.8	76.2	754	17	A415322
36	19.8	76.2	754	17	A415322
37	19.8	76.2	754	17	A415322
38	19.8	76.2	754	17	A415322
39	19.8	76.2	754	17	A415322
40	19.8	76.2	754	17	A415322
41	19.8	76.2	754	17	A415322
42	19.8	76.2	754	17	A415322
43	19.8	76.2	754	17	A415322
44	19.8	76.2	754	17	A415322
45	19.8	76.2	754	17	A415322

#### ALIGNMENTS

RESULT 1  
LOCUS: BH032631  
DEFINITION: BH032631.1 TV PFC1-24 Mus musculus genome clone PFC1-24.239N18  
ACCESSION: BH032631.1  
VERSION: G1:14803637  
KEYWORDS: GSS.  
SOURCE: house mouse.  
ORGANISM: Mus musculus.  
REFERENCE: Zhao, S., Nierman, W., Malek, J., Shatsman, S., Akint, R., Levins, M., Russell, D., de Jong, P., and Fraser, C. M. Mouse BAC End Sequences from Library PFC1-24  
Other GSS: PFC1-24.239N18.TJ  
COMMENT: Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: shatsman@igf.org  
Clones are derived from the mouse BAC library PFC1-24. For BAC library availability, please contact Pieter de Jong (pdejong@mail.nih.gov). Clones may be purchased from BACMAN.







VERSION AM529446.1 GI:7166341  
 KEYWORDS EST.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 popayota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 535)  
 REFERENCE Honaldo, M.F., Lennon, G. and Soares, M.B.  
 AUTHORS Normalization and subtraction: two approaches to facilitate gene  
 TITLE discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 COMMENT Contact: Soares, MB  
 Program for Rat Gene Discovery and Mapping  
 University of Iowa  
 451 Hirsteinn Medical Research Building Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: mssoares@blue.weeg.uiowa.edu  
 The sequence contained an oligo-dT track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a bonafide poly A  
 tail. The sequence tag present in the cDNA between the NciI site  
 and the oligo-dT track served to identify it as a clone from the  
 normalized hippocampus library cDNA library preparation: M.B.  
 Soares lab clone distribution: clones will be available through  
 Research Genetics (www.resgen.com) The following repetitive  
 elements were found in this cDNA sequence: 1-50,  
 >AT-richLow-complexity  
 Seq primer: M13 forward  
 POLYA-Yes.

FEATURES  
 source  
 location/Qualifiers  
 1..535  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="01-R-B00-ale-q-11-0-01"  
 /clone\_1lb="01-R-B00"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pTZ19-Pac (Pharmacia) with a modified  
 polylinker. Site\_1: Not I; Site\_2: Eco RI; The library  
 (01-R-B00) is a subtracted library derived from a mixture  
 of the following tissues: thalamus, cerebellum,  
 hypothalamus, medulla, pons, midbrain, cerebral cortex,  
 corpus striatum and hippocampus. For a detailed  
 description of the library from which this clone was  
 derived, please visit our web site at  
 ratseq.uiowa.edu. The subtraction has been previously  
 described in (Honaldo, Lennon and Soares, Genome Research  
 6:791-806, 1996)  
 TAG\_LIB=01-R-B00  
 TAG\_TISSUE=hippocampus  
 TAG\_SEQ=CATTC\*

BASE COUNT 196 a 107 c 62 g 175 t  
 ORIGIN

Query Match 76.2%; Score 19.8; DB 10; Length 535;  
 Best Local Similarity 91.3%; Pred. No. 3 aa+02;  
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 CTCGCCAATCAATTCATTTC 24  
 |||||||  
 DB 330 CTCGCCAATCAATTCATTTC 352

RESULT 9  
 LOCUS BE109206 577 bp mRNA linear EST 13-JUN-2000  
 DEFINITION U1-R-B51-azd-b-07-0-01.s1 U1-R-B51: Rattus norvegicus cDNA clone  
 SOURCE U1-R-B51-azd-b-07-0-01 3', mRNA sequence.  
 ACCESSION BE109206

VERSION BE109206.1 GI:8501311  
 KEYWORDS EST.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 popayota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 577)  
 REFERENCE Honaldo, M.F., Lennon, G. and Soares, M.B.  
 AUTHORS Normalization and subtraction: two approaches to facilitate gene  
 TITLE discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 COMMENT Contact: Soares, MB  
 Program for Rat Gene Discovery and Mapping  
 University of Iowa  
 451 Hirsteinn Medical Research Building Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: mssoares@blue.weeg.uiowa.edu  
 The sequence contained an oligo-dT track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a bonafide poly A  
 tail. The sequence tag present in the cDNA between the NciI site  
 and the oligo-dT track served to verify it as a clone from the  
 normalized embryo at 13 dpc library cDNA library preparation: M.B.  
 Soares lab clone distribution: clones will be available through  
 Research Genetics (www.resgen.com) The following repetitive  
 elements were found in this cDNA sequence: 1-50,  
 >AT-richLow-complexity  
 Seq primer: M13 forward  
 POLYA-Yes.

FEATURES  
 source  
 location/Qualifiers  
 1..577  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="01-R-B51-azd-b-07-0-01"  
 /clone\_1lb="01-R-B51"  
 /dev\_stage="embryonic 13 dpc"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pTZ19-Pac (Pharmacia) with a modified  
 polylinker. Site\_1: Not I; Site\_2: Eco RI; The 01-R-B51  
 library is derived from 13 dpc whole embryo tissue. For a  
 detailed description of the library from which this clone  
 was derived, please visit our web site at  
 ratseq.uiowa.edu.  
 TAG\_LIB=01-R-B51  
 TAG\_TISSUE=embryo at 13 dpc  
 TAG\_SEQ=AATTC\*

BASE COUNT 203 a 113 c 77 g 184 t  
 ORIGIN

Query Match 76.2%; Score 19.8; DB 10; Length 577;  
 Best Local Similarity 91.3%; Pred. No. 4a+02;  
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 CTCGCCAATCAATTCATTTC 24  
 |||||||  
 DB 330 CTCGCCAATCAATTCATTTC 352

RESULT 10  
 LOCUS BH093286 587 bp DNA linear GSS 18-JUL-2001  
 DEFINITION RPCI-24-297H24-TV RPCI-24 Mus musculus genomic clone RPCI-24-297H24  
 DNA sequence.  
 ACCESSION BH093286  
 LOCUS BH093286 GSS.  
 VERSION BH093286.1 GI:14913191  
 KEYWORDS Mus musculus.  
 SOURCE Mus musculus.  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



```

/ld_xref="taxon:10090"
/clone_lib="U9C1M0139C15"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-
/notes: Vector, pMD42uv, purified genomic DNA from M.
musculus C57Bl/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/shares/) The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared RNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (314732119b/AF:29572.1), a copy number
inducible derivative of plasmid M1. The vector was ligated
with adaptors complementary to the insect adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT      229 a      135 c      109 g      166 t
ORIGIN
Query Match      76.2%, Score 19.8; DB 17; Length 619;
Best Local Similarity 91.3%; Pred. No. 4,1e+02;
Matches 21, Conservative 0, Mismatches 2, Indels 0; Gaps 0;
OY      3 TCCCAATCAATTACATTTTCT 25
      11111111111111111111
Db      404 TCCCAATCAATTACATATCT 426
RESULT 13
LOCUS      AM331729 644 bp  linear  EST 31-JAN-2000
DEFINITION      614096C01 x1 614 - root cDNA library from Walbot lab Zea mays cDNA
      mRNA sequence.
ACCESSION      AM331729
VERSION      AM331729.1
KEYWORDS      EST.
SOURCE      Zea mays.
ORGANISM      Zea mays.
REFERENCE      Enkaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta,
      Spermatophyta; Magnoliophyta, Liliopsida, Poales, Poaceae; PACC
      clade; Panicoidae; Andropogoneae, Zea.
      1 (bases 1 to 644)
TITLE      Maize ESTs from various cDNA libraries sequenced at Stanford
      University
JOURNAL      Unpublished (1999)
COMMENT      Contact: Walbot V
      Department of Biological Sciences
      Stanford University
      855 California Ave, Palo Alto, CA 94304, USA
      Tel: 650 723 2927
      Fax: 650 725 6221
      Email: walbot@stanford.edu
      Plate: 614096 row: C column: 01.
FEATURES
      Location/Qualifiers
          1..644
             /organism="Zea mays"
             /cultivar="W23"
             /db_xref="taxon:4577"
             /clone_lib="614 - root cDNA library from Walbot lab"
             /issue_type="root"
             /dev_stage="3-4 days old"
             /lab_host="XL0LR"
             /note="Organ: root; Vector: pMutectScriptII SK1; Site: 1;
             Root: Site 2; XhoI: 3 4 days old root tissue from Walbot

```

BASE COUNT	177 a	123 c	126 g	217 t	1 others
ORIGIN					
Query Match		76.2%	Score 19.8;	DB 10;	Length 644;
Best Local Similarity	91.3%	Preq. No. 4,26,02;			
Matches	21;	Conservative	0;	Mismatches	2;
				Indels	0;
				Gaps	0;
27	3	TCCTCAATGATACATTTCTT	25		
Db	314	TCCTCAATGATACATTTCTT	252		
RESULT 14					
AZ873765		670 bp	DNA	linear	GSS 21-FEB-2001
LOCUS					
DEFINITION					
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
COMMENT					

BASE COUNT 241 a 143 c 108 g 188 t  
ORIGIN

Query Match  
Best local similarity 76.2%; Score 19.8; DB 17; Length 670;  
Best local similarity 91.3%; Pred. No. 4,4002;  
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

UY 3 CTGCGCAATCAATACATTTC 25  
|||||  
DB 493 CTGCGCAATCAATACATTTC 515

## RESULT 15

BO199536

DEFINITION M199536.683 bp. mRNA. Linear. EST 02-MAY-2002

UI R ERO c1c-b-08-0 01 s1 01 R ERO Rattus norvegicus cDNA clone

UI R ERO c1c-b-08-0 01 3', mRNA sequence.

BO199536

BO199536.1 61:20415991

EST

NOTWAY RAT

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus;

1 (bases 1 to 683)

Normalized and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 E. East Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: msoares@iuii.wiwi.iowa.edu

The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. The sequence lay present in the cDNA between the Molt site  
and the oligo-dT track served to verify it as a clone from the  
non-normalized duodenal library cDNA library preparation: M.H.  
Soares Lab clone distribution: clones will be available through  
Research Genetics (www.resgen.com) The following repetitive  
elements were found in this cDNA sequence: 1-51.  
-AT-rich low complexity  
Seq primer: M13 Forward  
polA-Tes.

Location/Qualifiers

1..683

Organism: "Rattus norvegicus"

Strain: "Sprague-Dawley"

DB\_Xref: "taxon:10116"

Clone: "01 R ERO c1c-b-08-0-01"

Clone: "10-01 R ERO"

Dev stage: "adult"

Lab host: "DH10B (Life Technologies)"

Note: Vector: PT733-Pac (Pharmacia) with a modified

polylinker. Site 1: Not 1; Site 2: Pro P1; nt-ERO is a

non-normalized Rat duodenal library (R0) constructed in

PT417 PAC vector according to the procedure described by

Honaido, Lennon & Soares (Genome Research Genome 6:

791-806, 1996). The oligonucleotide used to prime first

strand synthesis contained the sequence tag TGTGCTCAT

between the Not I cloning site and d18 stretch. The Rat

duodenal tissue was provided by Tom Freeman of the Sanger

Center.

TAG\_L1R-01 R ERO

TAG\_TISSUE: duodenum

TAG\_SEQ: TGTGCTCAT

BASE COUNT 229 a 138 c 100 g 216 t  
ORIGIN

Query Match  
Best local similarity 76.2%; Score 19.8; DB 14; Length 683;  
Best local similarity 91.3%; Pred. No. 4,4002;  
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

UY 2 CTGCGCAATCAATACATTTC 24  
|||||  
DB 331 CTGCGCAATCAATACATTTC 453

Search completed: January 14, 2003, 15:10:31  
Job time: 41.817 secs